JOURNAL OF
COMPUTATIONAL AND
APPLIED MATHEMATICS

ELSEVIER Journal of Computational and Applied Mathematics 91 (1998) 137–158

A posteriori and constructive a priori error bounds for finite element solutions of the Stokes equations

Mitsuhiro T. Nakao[a,*], Nobito Yamamoto[a], Yoshitaka Watanabe[b]

[a]*Graduate School of Mathematics, Kyushu University 33, Fukuoka 812-81, Japan*

[b]*Computer Center, Kyushu University 33, Fukuoka 812-81, Japan*

Received 28 August 1996; received in revised form 4 February 1998

Abstract

We describe a method to estimate the guaranteed error bounds of the finite element solutions for the Stokes problem in mathematically rigorous sense. We show that an a posteriori error can be computed by using the numerical estimates of a constant related to the so-called inf-sup condition for the continuous problem. Also a method to derive the constructive a priori error bounds are considered. Some numerical examples which confirm us the expected rate of convergence are presented. © 1998 Elsevier Science B.V. All rights reserved.

Keywords: Stokes equations; Guaranteed error bounds; Computable error

1. Introduction

In the numerical approximation of partial differential equations, it is very important to estimate the computable error bounds. For the finite element solutions of the Stokes equations, several a posteriori error bounds have been derived, e.g., [1, 2, 9, 11, 12]. In [13], Verfürth also applied his general results to finite element approximations of scalar quasi-linear elliptic partial differential equations of 2nd order and the stationary incompressible Navier–Stokes equations. However, these methods provide only information about the local and global error quantities of the computed finite element solutions controlling self-adaptive mesh-refinement processes, so they do not guarantee mathematically rigorous error bounds at all.

In this paper, we describe a method to estimate the guaranteed accuracy of the finite element solutions for the Stokes problem.

Using the numerical estimates of a constant related to the so-called inf-sup condition, we show an a posteriori error bound of the Stokes problem. Also we describe a method to derive the constructive a priori error bounds based on the estimation of the largest eigen value for matrices.

We emphasize that these results provide a basis of the numerical verification method (cf. [7, 14–16]) of the solution for the stationary Navier–Stokes equations.

PII S0377-0427(98)00035-1

1.1. The Stokes equations

Consider the following Stokes problem

$$\begin{aligned} -\nu\Delta u + \nabla p = f \quad & \text{in} \quad \Omega, \\ \operatorname{div} u = 0 \quad & \text{in} \quad \Omega, \\ u = 0 \quad & \text{on} \quad \partial\Omega, \end{aligned} \tag{1.1}$$

where $\nu > 0$ is the viscosity constant, $u = (u_1, u_2)^{\mathrm{T}}$ the two-dimensional velocity field, $f = (f_1, f_2)^{\mathrm{T}}$ a smooth function which means a density of body forces per unit mass and Ω a convex polygonal domain in $\mathbb{R}^2$. Here, p represents a kinematic pressure field and $\operatorname{div} u = 0$ means the incompressibility condition. Stokes equations (1.1) are the linearized stationary form of the Navier-Stokes equations.

An outline of this paper is as follows. Section 2 is devoted to derive some inequalities for the functions corresponding to velocity and pressure. It is based on the numerical estimates of a constant related to the so-called inf-sup condition for the continuous problem. In Section 3, we define the approximate solutions for the Stokes equations by using some finite element subspaces and show an a posteriori error bound under the suitable assumption. We propose a method to derive the constructive a priori error bounds and give a detailed computing algorithm in Section 4. And some numerical examples are presented in Section 5.

1.2. Some function spaces

We denote by $H^k(\Omega)$ the usual k-th order Sobolev space on Ω, and define $(\cdot,\cdot)$ as the inner product in $L^2(\Omega)$ and put

$$\begin{aligned} & H_0^1(\Omega) \equiv \{ v \in H^1(\Omega);\ v = 0 \ \text{ on } \ \partial\Omega \}, \\ & L_0^2(\Omega) \equiv \{ v \in L^2(\Omega);\ (v, 1) = 0 \}, \\ & \mathcal{S} \equiv H_0^1(\Omega)^2 \times L_0^2(\Omega). \end{aligned}$$

The norm in $L^2(\Omega)$ and $H_0^1(\Omega)$ are denoted by $|q|_0 \equiv (q,q)^{1/2}$, $|v|_1 \equiv |\nabla v|_0$, respectively. In what follows, since no confusion may arise, we will use the same notations for the corresponding norms and inner products in $L^2(\Omega)^2$ and $H_0^1(\Omega)^2$ as in $L^2(\Omega)$ and $H_0^1(\Omega)$, respectively.

We also define $H^2(\Omega)$-seminorm $|\cdot|_2$ by

$$|u|_2 \equiv \left(\left| \frac{\partial^2 u}{\partial x^2} \right|_0^2 + 2 \left| \frac{\partial^2 u}{\partial x \partial y} \right|_0^2 + \left| \frac{\partial^2 u}{\partial y^2} \right|_0^2 \right)^{1/2},$$

and set the function spaces

$$\begin{aligned} & V \equiv \{ v \in H_0^1(\Omega)^2;\ \operatorname{div} v = 0 \}, \\ & V^\perp \equiv \{ v \in H_0^1(\Omega)^2;\ (\nabla v \nabla w) = 0,\ w \in V \}. \end{aligned}$$

Finally, we define $H^{-1}(\Omega)^2$ as the dual space of $H_0^1(\Omega)^2$ and $\langle \cdot,\cdot \rangle$ as the duality pairing between $H^{-1}(\Omega)^2$ and $H_0^1(\Omega)^2$.

2. Numerical estimates for inf-sup condition

In this section, we rewrite the Stokes equations (1.1) to the weak form and give the numerical estimates of a constant related to the inf-sup condition which assures the existence of a weak solution. Using this constant, we present some norm inequalities for each element of $\mathscr{S}$.

2.1. Variational formulation

We introduce a bilinear form $\mathscr{L}$ on $\mathscr{S} \times \mathscr{S}$ by

$$\mathscr{L}([u,p],[v,q]) \equiv \nu(\nabla u, \nabla v) - (p, \operatorname{div} v) - (q, \operatorname{div} u), \quad [u,p],\ [v,q] \in \mathscr{S}. \tag{2.1}$$

Then, the standard variational formulation of (1.1) is given by

$$\begin{aligned} &\text{find } [u,p] \in \mathscr{S} \text{ such that} \\ &\mathscr{L}([u,p],[v,q]) = (f,v), \quad \forall [v,q] \in \mathscr{S}. \end{aligned} \tag{2.2}$$

It is well-known (cf. [4]) that (2.2) has a unique solution in $\mathscr{S}$ and there exists a constant $\beta_c > 0$ (depends only on Ω) such that $\mathscr{L}$ satisfies the following condition:

$$\inf_{\substack{[u,p]\in\mathscr{S} \\ [u,p]\neq 0}} \sup_{\substack{[v,q]\in\mathscr{S} \\ [v,q]\neq 0}} \frac{\mathscr{L}([u,p],[v,q])}{(|u|_1 + |p|_0)(|v|_1 + |q|_0)} \geqslant \beta_c. \tag{2.3}$$

From (2.3), we get

$$\sup_{\substack{[v,q]\in\mathscr{S} \\ [v,q]\neq 0}} \frac{\mathscr{L}([u,p],[v,q])}{|v|_1 + |q|_0} \geqslant \beta_c(|u|_1 + |p|_0), \quad \forall [u,p] \in \mathscr{S}.$$

Now, for $[u,p] \in \mathscr{S}$, we define $\delta(u,p)$ by

$$\delta(u,p) \equiv \sup_{\substack{[v,q]\in\mathscr{S} \\ [v,q]\neq 0}} \frac{\mathscr{L}([u,p],[v,q])}{|v|_1 + |q|_0}. \tag{2.4}$$

In the remainder of this section, we attempt to bound estimate $|u|_1$ and $|p|_0$ using above $\delta(u,p)$. For this purpose, we use the following lemma (proof is in [4]).

Lemma 2.1 (Babuška–Aziz). *For all $q \in L_0^2(\Omega)$, there exists a unique $v \in V^\perp$ such that*

$$\operatorname{div} v = q, \tag{2.5}$$

$$|v|_1 \leqslant \frac{1}{\beta} |q|_0. \tag{2.6}$$

where $\beta > 0$ is a constant depending only on Ω.

By virtue of Lemma 2.1, we obtain

$$\inf_{\substack{p\in L_0^2(\Omega) \\ p\neq 0}} \sup_{\substack{u\in H_0^1(\Omega)^2 \\ u\neq 0}} \frac{-(p, \operatorname{div} u)}{|u|_1 |p|_0} \geqslant \beta. \tag{2.7}$$

This condition (2.7) is called *inf-sup condition* for $\mathscr{L}$ which assures that problem (2.2) has a unique solution in $\mathscr{S}$. If Ω is a general bounded connected domain, it is difficult to estimate an explicit upper or lower bound of β. However, if Ω is the star shaped domains (i.e. including the case of convex polygons), this constant β can be numerically determined due to the following Horgan's lemma [5].

Lemma 2.2 (Horgan [5]). *Suppose that Ω is a star-shaped domain with respect to a point, which we choose to be the origin. Let the boundary be represented in plane polar coordinates by*

$$r = f(\theta) \quad \text{on} \quad \partial\Omega$$

and let

$$\mathscr{F}(\theta) \equiv \left\{ \left[1 + \left(\frac{f'(\theta)}{f(\theta)} \right)^2 \right]^{1/2} + \frac{|f'(\theta)|}{f(\theta)} \right\}^2.$$

Then, for the constant β in Lemma 2.1, the following estimate holds:

$$\frac{1}{\beta} \leqslant \sqrt{1 + \max_{\theta} \mathscr{F}(\theta)}. \tag{2.8}$$

For the special case, if Ω is a square, we have $1/\beta < 2.614$. Moreover, Horgan conjectures that the optimal constant for a square would be $1/\beta = \sqrt{7/2} \sim 1.871$.

2.2. Norm inequalities

Now, using a constant β, we can describe the following inequalities.

Theorem 2.3. *For all $[u, p] \in \mathscr{S}$, let us define $\delta(u, p)$ by (2.4), then the following estimates hold:*

$$\begin{aligned} |u|_1 &\leqslant \left(\frac{1}{\nu^2} + \frac{1}{\beta^2} \right)^{1/2} \delta(u, p), \\ |p|_0 &\leqslant \left(\frac{1}{\beta} + \frac{\nu}{\beta^2} \right) \delta(u, p). \end{aligned} \tag{2.9}$$

Proof. Since V is a closed subspace of $H_0^1(\Omega)^2$, we have the decomposition

$$H_0^1(\Omega)^2 = V \oplus V^{\perp}.$$

Therefore, each $u \in H_0^1(\Omega)^2$ can be written as

$$u = w + u_0 \quad w \in V, \quad u_0 \in V^{\perp}.$$

If $w \neq 0$, taking $v = w$, $q = 0$ in the definition of $\delta(u, p)$, we have

$$\delta(u, p) \geqslant \frac{\nu(\nabla w, \nabla w) - (p, \operatorname{div} w)}{|w|_1}.$$

Since $w \in V$, we get

$$\delta(u,p) \geqslant \nu|w|_1, \tag{2.10}$$

and this also holds for the case that $w=0$.

On the other hand, if $u_0 \neq 0$, taking $v=0$ in (2.4) implies

$$\delta(u,p)|q|_0 \geqslant -(q, \operatorname{div} u_0) \quad \forall q \in L_0^2(\Omega).$$

By Lemma 2.1, we can take q as $\operatorname{div} u_0 = -q$ and $\beta|u_0|_1 \leqslant |q|_0$. Then we have

$$\delta(u,p) \geqslant \beta|u_0|_1, \tag{2.11}$$

and it is clear that (2.11) holds for the case that $u_0=0$. Therefore, from (2.10) and (2.11) we have

$$\begin{aligned} |u|_1^2 &= |w|_1^2 + |u_0|_1^2 \\ &\leqslant \left(\frac{1}{\nu^2}+\frac{1}{\beta^2}\right)\delta(u,p)^2. \end{aligned}$$

Next, for all $p \in L_0^2(\Omega)$, if $p \neq 0$, by Lemma 2.1, we can take $v \in V^\perp$ satisfying

$$\operatorname{div} v = -p, \quad |v|_1 \leqslant \frac{1}{\beta}|p|_0.$$

Since $(\nabla u, \nabla v) = (\nabla u_0, \nabla v)$, we get

$$\delta(u,p) \geqslant \frac{\nu(\nabla u_0, \nabla v) - (q, \operatorname{div} u_0) + |p|_0^2}{|v|_1 + |q|_0} \quad \forall q \in L_0^2(\Omega).$$

If $u_0 \neq 0$ we take $q \in L_0^2(\Omega)$ such that

$$K = \frac{\nu(\nabla u_0, \nabla v)}{|\operatorname{div} u_0|_0^2}, \quad q = K \operatorname{div} u_0.$$

This implies that $\nu(\nabla u_0, \nabla v) - (q, \operatorname{div} u_0) = 0$. Moreover, from (2.6),

$$|u_0|_1 \leqslant \frac{1}{\beta}|\operatorname{div} u_0|_0.$$

Hence, we get

$$\begin{aligned} |q|_0 &= \frac{\nu(\nabla u_0, \nabla v)}{|\operatorname{div} u_0|_0} \\ &\leqslant \frac{\nu|u_0|_1|v|_1}{|\operatorname{div} u_0|_0} \\ &\leqslant \frac{\nu}{\beta^2}|p|_0, \end{aligned}$$

and this inequality follows for the case that $u_0=0$, by taking $q=0$.

Consequently, we obtain

$$\delta(u,p) \geqslant \frac{|p|_0^2}{\dfrac{1}{\beta}|p|_0 + \dfrac{\nu}{\beta^2}|p|_0} = \left(\frac{1}{\beta} + \frac{\nu}{\beta^2}\right)^{-1} |p|_0.$$

It is clear that the above inequality holds for $p = 0$. □

3. An a posteriori error bound

In this section, we introduce some finite element subspaces for the approximation of the velocity and the pressure, and show an a posteriori error bound for the Stokes equations using Theorem 2.3.

3.1. Finite element subspace

Let $\mathscr{T}_h$ be a family of triangulations of $\Omega \subset \mathbb{R}^2$, which consist of triangles or quadrilaterals dependent on a scale parameter $h > 0$. For $\mathscr{T}_h$, we denote by $X_h \subset H_0^1(\Omega) \cap C(\bar{\Omega})$ and $Y_h \subset L_0^2(\Omega) \cap C(\bar{\Omega})$ the finite element subspaces for the approximation of the velocity u and the pressure p, respectively. And we set $S_h \equiv X_h^2$. Then the standard finite element solution to (2.2) is defined by

$$\begin{aligned} &\text{find} \quad [u_h, p_h] \in S_h \times Y_h \quad \text{such that} \\ &\mathscr{L}([u_h, p_h], [v_h, q_h]) = (f, v_h), \quad \forall [v_h, q_h] \in S_h \times Y_h. \end{aligned} \tag{3.1}$$

Now, we introduce a post-processing procedures proposed by [15]. We define X_h^* as a subspace of $H^1(\Omega)$ in which the basis of X_h^* are the union of the basis of S_h and base functions corresponding to nodes on the boundary $\partial\Omega$. Note that

$$X_h \subset X_h^* \subset H^1(\Omega), \quad X_h \neq X_h^*.$$

We also define L^2-projection $P_0 : L^2(\Omega) \longrightarrow X_h$, L^2-projection $\hat{P}_0 : L^2(\Omega) \longrightarrow X_h^*$ and H_0^1-projection $P_1 : H_0^1(\Omega) \longrightarrow X_h$ by

$$\begin{aligned} &(v - P_0 v, \phi) = 0, && \forall \phi \in X_h, \\ &(v - \hat{P}_0 v, \phi) = 0, && \forall \phi \in X_h^*, \\ &(\nabla(v - P_1 v), \nabla\phi) = 0, && \forall \phi \in X_h, \end{aligned}$$

respectively. For $w_h \in X_h$, we define $\overline{\nabla} w_h \in (X_h^*)^2$ and $\overline{\Delta} w_h \in L^2(\Omega)$ by

$$\begin{aligned} &\overline{\nabla} w_h \equiv \left(\hat{P}_0 \frac{\partial w_h}{\partial x}, \hat{P}_0 \frac{\partial w_h}{\partial y}\right), \\ &\overline{\Delta} w_h \equiv \operatorname{div} \overline{\nabla} w_h, \end{aligned}$$

respectively. By virtue of a direct consequence of the argument in [15], it can be shown that for all $v_h \in S_h$, the following properties hold:

$$(-\overline{\Delta} v_h, \phi) = (\overline{\nabla} v_h, \nabla \phi) \quad \forall \phi \in H_0^1(\Omega)^2, \tag{3.2}$$

$$|\overline{\nabla} v_h - \nabla v_h|_0 = \inf_{w_h \in (X_h^*)^2 \times (X_h^*)^2} |w_h - \nabla v_h|_0. \tag{3.3}$$

Now, we assume, as the approximation property of X_h the following.

Assumption 3.1

$$\inf_{\xi \in X_h} |v - \xi|_1 \leqslant C_0 h |v|_2, \quad \forall v \in H_0^1(\Omega) \cap H^2(\Omega), \tag{3.4}$$

where C_0 is a positive constant independent of v and h which can be numerically determined.

This assumption holds for many finite element subspaces (cf.[3, 6, 8]). From the properties of projection P_1, Assumption 3.1 and Aubin–Nitche's trick, for all $v \in H_0^1(\Omega)$,

$$|v - P_1 v|_1 \leqslant |v|_1, \tag{3.5}$$

$$|v - P_1 v|_0 \leqslant C_0 h |v|_1 \tag{3.6}$$

hold.

3.2. An a posteriori estimate

Let $[u, p]$ and $[u_h, p_h]$ be the solutions of (2.2) and (3.1), respectively. And set

$$e_h = u - u_h,$$
$$\varepsilon_h = p - p_h.$$

We will now estimate an upper bound of $\delta(e_h, \varepsilon_h)$. For all $[v, q] \in \mathscr{S}$, from the definition of $\mathscr{L}$, we get

$$\mathscr{L}([e_h, \varepsilon_h], [v, q]) = \nu(\nabla e_h, \nabla v) - (\varepsilon_h, \operatorname{div} v) - (q, \operatorname{div} e_h). \tag{3.7}$$

On the other hand, by (2.2) and (3.1), for all $[\xi_h, q_h] \in S_h \times Y_h$,

$$\nu(\nabla e_h, \nabla \xi_h) - (\varepsilon_h, \operatorname{div} \xi_h) - (q_h, \operatorname{div} e_h) = 0 \tag{3.8}$$

holds. Taking $q_h = 0$ in (3.8), for all $\xi_h \in S_h$, we have

$$\nu(\nabla e_h, \nabla \xi_h) - (\varepsilon_h, \operatorname{div} \xi_h) = 0. \tag{3.9}$$

From (3.7) and (3.9) we have

$$\begin{aligned}\mathscr{L}([e_h, \varepsilon], [v, q]) &= \nu(\nabla e_h, \nabla(v - \xi_h)) - (\varepsilon_h, \operatorname{div}(v - \xi_h)) - (q, \operatorname{div} e_h) \\ &= \nu(\nabla(u - u_h), \nabla(v - \xi_h)) - (p - p_h, \operatorname{div}(v - \xi_h)) + (q, \operatorname{div} u_h).\end{aligned}$$

Moreover, by virtue of (3.2), Green's formula and Schwarz's inequality, we obtain

$$\begin{aligned}
\mathscr{L}([e_h,\varepsilon_h],[v,q]) &= v(\overline{\nabla} u_h - \nabla u_h, \nabla(v-\xi_h)) + v(\nabla u - \overline{\nabla} u_h, \nabla(v-\xi_h)) \\
&\quad -(p-p_h, \operatorname{div}(v-\xi_h)) + (q, \operatorname{div} u_h) \\
&= v(\overline{\nabla} u_h - \nabla u_h, \nabla(v-\xi_h)) + v\langle -\Delta u + \overline{\Delta} u_h, v-\xi_h\rangle \\
&\quad +\langle \nabla(p-p_h), v-\xi_h\rangle + (q, \operatorname{div} u_h) \\
&= v(\overline{\nabla} u_h - \nabla u_h, \nabla(v-\xi_h)) + (f + v\overline{\Delta} u_h - \nabla p_h, v-\xi_h) + (q, \operatorname{div} u_h) \\
&\leqslant v|\overline{\nabla} u_h - \nabla u_h|_0 \; |v-\xi_h|_1 + |v\overline{\Delta} u_h - \nabla p_h + f|_0 |v-\xi_h|_0 + |\operatorname{div} u_h|_0 \; |q|_0.
\end{aligned}$$

Now, we set $\xi_h \in S_h$ as an elementwise H_0^1-projection of v, i.e., $\xi_h = (P_1 v_1, P_1 v_2)$, then using (3.5) and (3.6), we have

$$\mathscr{L}([e_h,\varepsilon_h],[v,q]) \leqslant (v|\overline{\nabla} u_h - \nabla u_h|_0 + C_0 h|v\overline{\Delta} u_h - \nabla p_h + f|_0 + |\operatorname{div} u_h|_0)(|v|_1 + |q|_0).$$

Thus, the following result is obtained:

Lemma 3.2. *For all* $[v,q] \neq 0$ *in* $\mathscr{S}$,

$$\frac{\mathscr{L}([e_h,\varepsilon_h],[v,q])}{|v|_1 + |q|_0} \leqslant v|\overline{\nabla} u_h - \nabla u_h|_0 + C_0 h|v\overline{\Delta} u_h - \nabla p_h + f|_0 + |\operatorname{div} u_h|_0$$

holds.

From Theorem 2.3 and Lemma 3.2, we have the following a posteriori error bounds for finite element solutions of the Stokes equations.

Theorem 3.3 (A posteriori error bounds). *Let* $[u,p]$ *and* $[u_h,p_h]$ *be solutions of (2.2) and (3.1), respectively. Then, the following a posteriori error bounds are obtained:*

$$\begin{aligned}
|u-u_h|_1 &\leqslant \left(\frac{1}{v^2} + \frac{1}{\beta^2}\right)^{1/2} C(u_h,p_h), \\
|p-p_h|_0 &\leqslant \left(\frac{1}{\beta} + \frac{v}{\beta^2}\right) C(u_h,p_h),
\end{aligned} \tag{3.10}$$

where $C(u_h,p_h)$ *is an a posteriori error estimator which can be computed using the finite element solutions* $[u_h,p_h]$ *by*

$$C(u_h,p_h) \equiv v|\overline{\nabla} u_h - \nabla u_h|_0 + C_0 h|v\overline{\Delta} u_h - \nabla p_h + f|_0 + |\operatorname{div} u_h|_0. \tag{3.11}$$

Proof. From Theorem 2.3, we have immediately

$$\begin{aligned}
|u-u_h|_1 &\leqslant \left(\frac{1}{v^2} + \frac{1}{\beta^2}\right)^{1/2} \delta(e_h,\varepsilon_h), \\
|p-p_h|_0 &\leqslant \left(\frac{1}{\beta} + \frac{v}{\beta^2}\right) \delta(e_h,\varepsilon_h).
\end{aligned}$$

On the other hand, from Lemma 3.2,

$$\delta(e_h, \varepsilon_h) = \sup_{\substack{[v,q]\in\mathscr{V} \\ [v,q]\neq 0}} \frac{\mathscr{L}([e_h, \varepsilon_h], [v,q])}{|v|_1 + |q|_0} \leqslant C(u_h, p_h)$$

holds, then we get the desired conclusion. □

By virtue of (3.1) and (3.3), it is expected that each term in the right-hand side of (3.11) tends to be smaller as h is.

4. Constructive a priori error bounds

In Section 3, we proposed an a posteriori error bound for finite element solutions of the Stokes problem. In this section, using the similar techniques in the previous section, we consider two kinds of method to derive the constructive a priori error bounds and describe a computational procedure for the estimation of a priori constants.

4.1. A priori estimates

The first method is based on the result of an a posteriori error bound proposed in Section 3. For $f \in L^2(\Omega)^2$, we define $P_0 f \in S_h$ by

$$P_0 f = (P_0 f_1, P_0 f_2)^{\mathrm{T}}.$$

The property of L^2-projection implies

$$|f - P_0 f|_0^2 = |f|_0^2 - |P_0 f|_0^2,$$

hence, we can write for some $0 \leqslant \theta \leqslant \pi/2$,

$$\begin{aligned} |P_0 f|_0 &= |f|_0 \sin\theta, \\ |f - P_0 f|_0 &= |f|_0 \cos\theta. \end{aligned} \tag{4.1}$$

Now, we suppose that there exist the constants K_1, K_2 and K_3 such that

$$|\overline{\nabla} u_h - \nabla u_h|_0 \leqslant K_1 |P_0 f|_0, \tag{4.2}$$

$$|\overline{\Delta} u_h - \nabla p_h + P_0 f|_0 \leqslant K_2 |P_0 f|_0, \tag{4.3}$$

$$|\mathrm{div}\, u_h|_0 \leqslant K_3 |P_0 f|_0, \tag{4.4}$$

independent of $f \in L^2(\Omega)^2$. We describe later how to determine these constants. Then, we have the following theorem.

Theorem 4.1 (A priori error bound I). *For each* $f \in L^2(\Omega)^2$, *we have*

$$
\begin{aligned}
|u-u_h|_1 &\leqslant \left(\frac{1}{\nu^2}+\frac{1}{\beta^2}\right)^{1/2} C(h)\,|f|_0,\\
|p-p_h|_0 &\leqslant \left(\frac{1}{\beta}+\frac{\nu}{\beta^2}\right) C(h)\,|f|_0,
\end{aligned}
\tag{4.5}
$$

where

$$
C(h)=\sqrt{(\nu K_1+C_0hK_2+K_3)^2+(C_0h)^2}. \tag{4.6}
$$

Proof. For all $f \in L^2(\Omega)^2$, from (3.11) and (4.1)–(4.4), we have

$$
\begin{aligned}
C(u_h,p_h) &= \nu|\overline{\nabla}u_h-\nabla u_h|_0+C_0h|\nu\overline{\Delta}u_h-\nabla p_h+P_0f+f-P_0f|_0+|\operatorname{div}u_h|_0\\
&\leqslant \nu K_1|P_0f|_0+C_0h(K_2|P_0f|_0+|f-P_0f|_0)+K_3|P_0f|_0\\
&= ((\nu K_1+C_0hK_2+K_3)\sin\theta+C_0h\cos\theta)\,|f|_0\\
&\leqslant ((\nu K_1+C_0hK_2+K_3)^2+(C_0h)^2)^{1/2}\,|f|_0\\
&= C(h)\,|f|_0. \qquad \square
\end{aligned}
$$

The second method is the direct estimation without $\overline{\nabla}u_h$ or $\overline{\Delta}u_h$. From (3.7) and (3.9), we have, for each $\xi_h \in S_h$,

$$
\mathscr{L}([e_h,\varepsilon_h],[v,q])=\nu(\nabla(u-u_h),\nabla(v-\xi_h))-(p-p_h,\operatorname{div}(v-\xi_h))+(q,\operatorname{div}u_h).
$$

Now, taking ξ_h as

$$
v_h \equiv (P_1v_1,P_1v_2)^{\mathrm{T}},
$$

the property of H_0^1-projection implies that

$$
\mathscr{L}([e_h,\varepsilon_h],[v,q])=\nu(\nabla u,\nabla(v-v_h))-(p-p_h,\operatorname{div}(v-v_h))+(q,\operatorname{div}u_h).
$$

Hence, by Green's formula, Schwarz's inequality and (3.6), we get

$$
\begin{aligned}
\mathscr{L}([e_h,\varepsilon_h],[v,q]) &= (f-\nabla p_h,v-v_h)+(q,\operatorname{div}u_h)\\
&\leqslant |f-\nabla p_h|_0\,C_0h\,|v|_1+|q|_0\,|\operatorname{div}u_h|_0\\
&\leqslant (C_0h|f-\nabla p_h|_0+|\operatorname{div}u_h|_0)(|v|_1+|q|_0).
\end{aligned}
$$

Therefore, we have the following lemma.

Lemma 4.2. *For all* $0 \neq [v,q] \in \mathscr{S}$,

$$
\frac{\mathscr{L}([e_h,\varepsilon_h],[v,q])}{|v|_1+|q|_0} \leqslant C_0h|f-\nabla p_h|_0+|\operatorname{div}u_h|_0
$$

holds.

Therefore, by virtue of Lemma 4.2, we can also take $C(u_h, p_h)$ in Theorem 3.3 as

$$C(u_h, p_h) = C_0 h|f - \nabla p_h|_0 + |\operatorname{div} u_h|_0.$$

We now suppose that the constants K_4 can be taken as

$$|-\nabla p_h + P_0 f|_0 \leqslant K_4 |P_0 f|_0 \tag{4.7}$$

independent of $f \in L^2(\Omega)^2$. Here, K_4 will be determined in later part of this section. Then, we obtain another a priori error bound as follows:

Theorem 4.3 (A priori error bound II). *For all $f \in L^2(\Omega)^2$, it holds that*

$$\begin{aligned} |u - u_h|_1 &\leqslant \left(\frac{1}{\nu^2} + \frac{1}{\beta^2}\right)^{1/2} C(h)\,|f|_0, \\ |p - p_h|_0 &\leqslant \left(\frac{1}{\beta} + \frac{\nu}{\beta^2}\right) C(h)\,|f|_0, \end{aligned} \tag{4.8}$$

where

$$C(h) = \sqrt{(C_0 h K_4 + K_3)^2 + (C_0 h)^2}. \tag{4.9}$$

Proof. For all $f \in L^2(\Omega)^2$, from (4.4), (4.7) and (4.1), we obtain

$$\begin{aligned} C(u_h, p_h) &= C_0 h|-\nabla p_h + P_0 f + f - P_0 f|_0 + |\operatorname{div} u_h|_0 \\ &\leqslant C_0 h(K_4 |P_0 f|_0 + |f - P_0 f|_0) + K_3 |P_0 f|_0 \\ &= ((C_0 h K_4 + K_3)\sin\theta + C_0 h \cos\theta)\,|f|_0 \\ &\leqslant C(h)\,|f|_0. \qquad \square \end{aligned}$$

4.2. Computation of the constants $C(h)$

Now, we show a method to estimate a priori constant $C(h)$ that appeared in (4.6) and (4.9).

Let us define $\dim X_h = n$, $\dim Y_h = m$ and $\dim X_h^* = \hat{n}$. From the definition of X_h and X_h^*, we have $\hat{n} > n$. Next, we denote base functions of X_h, Y_h and X_h^* by $\{\phi_j\}_{1 \leqslant j \leqslant n}$, $\{\psi_j\}_{1 \leqslant j \leqslant m}$ and $\{\hat{\phi}_j\}_{1 \leqslant j \leqslant \hat{n}}$, respectively.

Now, using real coefficients $\{a_j^{(1)}\}_{1 \leqslant j \leqslant n}$, $\{a_j^{(2)}\}_{1 \leqslant j \leqslant n}$ and $\{b_j\}_{1 \leqslant j \leqslant m}$, we can uniquely represent the finite element solution $u_h = (u_h^{(1)}, u_h^{(2)})^{\mathrm{T}} \in S_h$ and $p_h \in Y_h$ of the form:

$$\begin{aligned} u_h^{(1)} &= \sum_{i=1}^{n} a_i^{(1)} \phi_i, \\ u_h^{(2)} &= \sum_{i=1}^{n} a_i^{(2)} \phi_i, \\ p_h &= \sum_{i=1}^{m} b_i \psi_i. \end{aligned}$$

Then, using base functions of X_h, Y_h, we rewrite (3.1) as

$$\begin{aligned}
&\sum_{i=1}^{n} a_i^{(1)}(\nabla\phi_i,\nabla\phi_j) - \sum_{i=1}^{n} b_i\left(\psi_i,\frac{\partial\phi_j}{\partial x}\right) = (f_1,\phi_j), \quad 1\leqslant j\leqslant n,\\
&\sum_{i=1}^{n} a_i^{(2)}(\nabla\phi_i,\nabla\phi_j) - \sum_{i=1}^{n} b_i\left(\psi_i,\frac{\partial\phi_j}{\partial y}\right) = (f_2,\phi_j), \quad 1\leqslant j\leqslant n,\\
&-\sum_{i=1}^{m} a_i^{(1)}\left(\psi_j,\frac{\partial\phi_i}{\partial x}\right) - \sum_{i=1}^{m} a_i^{(2)}\left(\psi_j,\frac{\partial\phi_i}{\partial y}\right) = 0, \quad 1\leqslant j\leqslant m.
\end{aligned}\tag{4.10}$$

Now, we define, denoting the $M\times N$ matrix by $(\cdots)_{M\times N}$,

$$\begin{aligned}
\boldsymbol{a}_1 &= (a_1^{(1)},a_2^{(1)},\ldots,a_n^{(1)})_{1\times n},\\
\boldsymbol{a}_2 &= (a_1^{(2)},a_2^{(2)},\ldots,a_n^{(2)})_{1\times n},\\
\boldsymbol{a} &= (\boldsymbol{a}_1,\boldsymbol{a}_2)_{1\times 2n},\\
\boldsymbol{b} &= (b_1,b_2,\ldots,b_m)_{1\times m},\\
\boldsymbol{f}_1 &= (\,(f_1,\phi_1),(f_1,\phi_2),\ldots,(f_1,\phi_n))^{\mathrm{T}}_{n\times 1},\\
\boldsymbol{f}_2 &= (\,(f_2,\phi_1),(f_2,\phi_2),\ldots,(f_2,\phi_n))^{\mathrm{T}}_{n\times 1},\\
\boldsymbol{f} &= \begin{pmatrix}\boldsymbol{f}_1\\ \boldsymbol{f}_2\end{pmatrix}_{2n\times 1}.
\end{aligned}$$

Additionally, we set

$$\begin{aligned}
(D_0)_{ij} &= (\nabla\phi_i,\nabla\phi_j)_{n\times n},\\
(D)_{ij} &= \begin{pmatrix} D_0 & 0\\ 0 & D_0\end{pmatrix}_{2n\times 2n},\\
(E_x)_{ij} &= \left(\psi_i,\frac{\partial\phi_j}{\partial x}\right) = -\left(\frac{\partial\psi_i}{\partial x},\phi_j\right)_{m\times n},\\
(E_y)_{ij} &= \left(\psi_i,\frac{\partial\phi_j}{\partial y}\right) = -\left(\frac{\partial\psi_i}{\partial y},\phi_j\right)_{m\times n},\\
(E)_{ij} &= \begin{pmatrix}E_x & E_y\end{pmatrix}_{m\times 2n},\\
(G)_{ij} &= \begin{pmatrix} D & -E^{\mathrm{T}}\\ -E & 0\end{pmatrix}_{(2n+m)\times(2n+m)}.
\end{aligned}$$

Then, (4.10) can be represented as

$$\begin{pmatrix} D_0 & 0 & -(E_x)^{\mathrm{T}} \\ 0 & D_0 & -(E_y)^{\mathrm{T}} \\ -E_x & -E_y & 0 \end{pmatrix} \begin{pmatrix} \boldsymbol{a}_1^{\mathrm{T}} \\ \boldsymbol{a}_2^{\mathrm{T}} \\ \boldsymbol{b}^{\mathrm{T}} \end{pmatrix} = \begin{pmatrix} \boldsymbol{f}_1 \\ \boldsymbol{f}_2 \\ 0 \end{pmatrix}$$

Consequently, (3.1) is equivalent to the following linear equation:

$$G \begin{pmatrix} \boldsymbol{a}^{\mathrm{T}} \\ \boldsymbol{b}^{\mathrm{T}} \end{pmatrix} = \begin{pmatrix} \boldsymbol{f} \\ 0 \end{pmatrix}. \tag{4.11}$$

We assume that the symmetric matrix G is invertible, which usually follows by so-called discrete inf-sup condition, and the inverse matrix G^{-1} is represented as

$$(G^{-1})_{ij} = \begin{pmatrix} G_a & G_b^{\mathrm{T}} \\ G_b & G_* \end{pmatrix}_{(2n+m)\times(2n+m)},$$

where G_a, G_b and G_* are $2n \times 2n$, $m \times 2n$, $m \times m$ matrices, respectively. Then, we have the following representation of the finite element solution $[u_h, p_h] \in S_h \times Y_h$ satisfying (3.1).

$$\begin{aligned} \boldsymbol{a}^{\mathrm{T}} &= G_a \boldsymbol{f}, \\ \boldsymbol{b}^{\mathrm{T}} &= G_b \boldsymbol{f}. \end{aligned} \tag{4.12}$$

Next, we represent $\overline{\nabla} u_h$ using coefficients of u_h. Since $\overline{\nabla} u_h = \left(\overline{\nabla} u_h^{(1)}, \overline{\nabla} u_h^{(2)}\right) \in (X_h^*)^2 \times (X_h^*)^2$, using real coefficients $\{c_j^{(1)}\}_{1 \leqslant j \leqslant \hat{n}}$, $\{c_j^{(2)}\}_{1 \leqslant j \leqslant \hat{n}}$, $\{d_j^{(1)}\}_{1 \leqslant j \leqslant \hat{n}}$ and $\{d_j^{(2)}\}_{1 \leqslant j \leqslant \hat{n}}$, we can uniquely write $\overline{\nabla} u_h$ in the form

$$\overline{\nabla} u_h^{(1)} = \left(\sum_{i=1}^{\hat{n}} c_i^{(1)} \hat{\phi}_i, \sum_{i=1}^{\hat{n}} d_i^{(1)} \hat{\phi}_i \right)^{\mathrm{T}},$$

$$\overline{\nabla} u_h^{(2)} = \left(\sum_{i=1}^{\hat{n}} c_i^{(2)} \hat{\phi}_i, \sum_{i=1}^{\hat{n}} d_i^{(2)} \hat{\phi}_i \right)^{\mathrm{T}}.$$

Also set

$$\begin{aligned} \boldsymbol{c}_1 &= (c_1^{(1)}, c_2^{(1)}, \cdots, c_{\hat{n}}^{(1)})_{1\times\hat{n}}, \quad & \boldsymbol{d}_1 &= (d_1^{(1)}, d_2^{(1)}, \cdots, d_{\hat{n}}^{(1)})_{1\times\hat{n}}, \\ \boldsymbol{c}_2 &= (c_1^{(2)}, c_2^{(2)}, \cdots, c_{\hat{n}}^{(2)})_{1\times\hat{n}}, \quad & \boldsymbol{d}_2 &= (d_1^{(2)}, d_2^{(2)}, \cdots, d_{\hat{n}}^{(2)})_{1\times\hat{n}}. \end{aligned}$$

From the definition of L^2-projection, for each $1 \leqslant j \leqslant \hat{n}$, we have

$$\sum_{i=1}^{\hat{n}} c_i^{(1)} \left(\hat{\phi}_i, \hat{\phi}_j\right) = \sum_{i=1}^{n} a_i^{(1)} \left(\frac{\partial \phi_i}{\partial x}, \hat{\phi}_j\right),$$

$$\sum_{i=1}^{\hat{n}} d_i^{(1)} \left(\hat{\phi}_i, \hat{\phi}_j\right) = \sum_{i=1}^{n} a_i^{(1)} \left(\frac{\partial \phi_i}{\partial y}, \hat{\phi}_j\right).$$

Therefore, setting

$$(\hat{L})_{ij} = \left(\hat{\phi}_i, \hat{\phi}_j\right)_{\hat{n}\times\hat{n}},$$
$$(K^x)_{ij} = \left(\frac{\partial \phi_i}{\partial x}, \hat{\phi}_j\right) = -\left(\phi_i, \frac{\partial \hat{\phi}_j}{\partial x}\right)_{n\times\hat{n}},$$
$$(K^y)_{ij} = \left(\frac{\partial \phi_i}{\partial y}, \hat{\phi}_j\right) = -\left(\phi_i, \frac{\partial \hat{\phi}_j}{\partial y}\right)_{n\times\hat{n}},$$

we get

$$\boldsymbol{c}_1\hat{L} = \boldsymbol{a}_1 K^x, \quad \boldsymbol{d}_1\hat{L} = \boldsymbol{a}_1 K^y.$$

Similarly, we have

$$\boldsymbol{c}_2\hat{L} = \boldsymbol{a}_2 K^x, \quad \boldsymbol{d}_2\hat{L} = \boldsymbol{a}_2 K^y.$$

Hence, the following relations hold because of the invertibility of $\hat{L}$.

Lemma 4.4. *If we set $n \times \hat{n}$ matrices $M^x = K^x\hat{L}^{-1}, M^y = K^y\hat{L}^{-1}$, then $\overline{\nabla} u_h$ is represented as*

$$\begin{aligned} \boldsymbol{c}_1 &= \boldsymbol{a}_1 M^x, \quad \boldsymbol{d}_1 = \boldsymbol{a}_1 M^y, \\ \boldsymbol{c}_2 &= \boldsymbol{a}_2 M^x, \quad \boldsymbol{d}_2 = \boldsymbol{a}_2 M^y. \end{aligned} \tag{4.13}$$

Now, we describe how to estimate $|P_0 f|_0^2$, $|\overline{\nabla} u_h - \nabla u_h|_0$, $|\nu\overline{\Delta} u_h - \nabla p_h + f|_0$, $|\operatorname{div} u_h|_0$ and $|-\nabla p_h + P_0 f|_0$ using the vector f.

Lemma 4.5. *We define the $n \times n$ matrix L and the $2n \times 2n$ matrix F by*

$$(L)_{ij} = (\phi_i, \phi_j)_{n\times n},$$
$$(F)_{ij} = \begin{pmatrix} L^{-1} & 0 \\ 0 & L^{-1} \end{pmatrix}_{2n\times 2n}.$$

Then, $|P_0 f|_0^2$ can be represented by a quadratic form:

$$|P_0 f|_0^2 = \boldsymbol{f}^{\mathrm{T}} F \boldsymbol{f}. \tag{4.14}$$

Proof. For each $f = (f_1, f_2)^{\mathrm{T}}$, using real coefficients $\{q_j^{(1)}\}_{1\leqslant j\leqslant n}$ and $\{q_j^{(2)}\}_{1\leqslant j\leqslant n}$ we can write

$$P_0 f_1 = \sum_{i=1}^{n} q_i^{(1)} \phi_i, \quad P_0 f_2 = \sum_{i=1}^{n} q_i^{(2)} \phi_i.$$

We set

$$\boldsymbol{q}_1 = (q_1^{(1)}, \ldots, q_n^{(1)})_{1\times n}, \quad \boldsymbol{q}_2 = (q_1^{(2)}, \ldots, q_n^{(2)})_{1\times n}.$$

Then, we have

$$q_1^{\mathrm{T}} = L^{-1}f_1, \quad q_2^{\mathrm{T}} = L^{-1}f_2.$$

Therefore, we obtain

$$\begin{aligned}|P_0 f|_0^2 &= (P_0 f_1, P_0 f_1) + (P_0 f_2, P_0 f_2)\\ &= q_1 L q_1^{\mathrm{T}} + q_2 L q_2^{\mathrm{T}}\\ &= f^{\mathrm{T}} F f. \qquad \square\end{aligned}$$

Lemma 4.6. *We define the* $2n \times 2n$ *matrix* Q_1, A_1 *by*

$$(Q_1)_{ij} = \begin{pmatrix} D_0 - M^x(K^x)^{\mathrm{T}} - M^y(K^y)^{\mathrm{T}} & 0 \\ 0 & D_0 - M^x(K^x)^{\mathrm{T}} - M^y(K^y)^{\mathrm{T}} \end{pmatrix}_{2n\times 2n},$$

$$(A_1)_{ij} = (\, G_a Q_1 G_a)_{2n\times 2n}.$$

Then, K_1 *can be estimated as follows:*

$$K_1 \leqslant \left(\sup_{x\in\mathbb{R}^{2n}} \frac{x^{\mathrm{T}} A_1\, x}{x^{\mathrm{T}} F\, x} \right)^{1/2}. \tag{4.15}$$

Proof. From the definition of L^2-projection and Lemma 4.4, we get

$$\begin{aligned}|\overline{\nabla} u_h - \nabla u_h|_0^2 &= (\nabla u_h, \nabla u_h) - (\overline{\nabla} u_h, \overline{\nabla} u_h)\\ &= a D a^{\mathrm{T}} - c_1 \hat{L} c_1^{\mathrm{T}} - d_1 \hat{L} d_1^{\mathrm{T}} - c_2 \hat{L} c_2^{\mathrm{T}} - d_2 \hat{L} d_2^{\mathrm{T}}\\ &= a D a^{\mathrm{T}} - a_1 (M^x(K^x)^{\mathrm{T}} + M^y(K^y)^{\mathrm{T}}) a_1^{\mathrm{T}} - a_2 (M^x(K^x)^{\mathrm{T}} + M^y(K^y)^{\mathrm{T}}) a_2^{\mathrm{T}}\\ &= (a_1, a_2) \begin{pmatrix} D_0 - M^x(K^x)^{\mathrm{T}} + M^y(K^y)^{\mathrm{T}} & 0 \\ 0 & D_0 - M^x(K^x)^{\mathrm{T}} + M^y(K^y)^{\mathrm{T}} \end{pmatrix} \begin{pmatrix} a_1^{\mathrm{T}} \\ a_2^{\mathrm{T}} \end{pmatrix}\\ &= a Q_1 a^{\mathrm{T}}\\ &= f^{\mathrm{T}} G_a Q_1 G_a f\\ &= f^{\mathrm{T}} A_1 f. \qquad \square\end{aligned}$$

Next, we define $\hat{n} \times \hat{n}$ matrices $\hat{D}^{xx}$, $\hat{D}^{xy}$, $\hat{D}^{yy}$ as

$$(\hat{D}^{xx})_{ij} = \left(\frac{\partial \hat{\phi}_i}{\partial x}, \frac{\partial \hat{\phi}_j}{\partial x} \right)_{\hat{n}\times\hat{n}}, \qquad (\hat{D}^{xy})_{ij} = \left(\frac{\partial \hat{\phi}_i}{\partial x}, \frac{\partial \hat{\phi}_j}{\partial y} \right)_{\hat{n}\times\hat{n}}, \qquad (\hat{D}^{yy})_{ij} = \left(\frac{\partial \hat{\phi}_i}{\partial y}, \frac{\partial \hat{\phi}_j}{\partial y} \right)_{\hat{n}\times\hat{n}},$$

and $n \times n$ matrices $\hat{E}^{xx}$, $\hat{E}^{xy}$ and $\hat{E}^{yy}$ as

$$(\hat{E}^{xx})_{ij} = M^x \hat{D}^{xx} (M^x)^{\mathrm{T}}_{n\times n}, \qquad (\hat{E}^{xy})_{ij} = M^x \hat{D}^{xy} (M^y)^{\mathrm{T}}_{n\times n}, \qquad (\hat{E}^{yy})_{ij} = M^y \hat{D}^{yy} (M^y)^{\mathrm{T}}_{n\times n},$$

and set $2n \times 2n$ matrix E_1 by

$$(E_1)_{ij} = \begin{pmatrix} \hat{E}^{xx} + \hat{E}^{xy} + (\hat{E}^{xy})^{\mathrm{T}} + \hat{E}^{yy} & 0 \\ 0 & \hat{E}^{xx} + \hat{E}^{xy} + (\hat{E}^{xy})^{\mathrm{T}} + \hat{E}^{yy} \end{pmatrix}_{2n\times 2n}.$$

We also define $m \times m$ matrix $\tilde{D}$ by

$$(\tilde{D})_{ij} = (\nabla\psi_i, \nabla\psi_j)_{m\times m}$$

and $\hat{n} \times m$ matrices $\hat{F}^{xx}$, $\hat{F}^{xy}$, $\hat{F}^{yx}$ and $\hat{F}^{yy}$ by

$$(\hat{F}^{xx})_{ij} = \left(\frac{\partial\hat{\phi}_i}{\partial x}, \frac{\partial\psi_j}{\partial x}\right)_{\hat{n}\times m}, \qquad (\hat{F}^{xy})_{ij} = \left(\frac{\partial\hat{\phi}_i}{\partial x}, \frac{\partial\psi_j}{\partial y}\right)_{\hat{n}\times m},$$

$$(\hat{F}^{yx})_{ij} = \left(\frac{\partial\hat{\phi}_i}{\partial y}, \frac{\partial\psi_j}{\partial x}\right)_{\hat{n}\times m}, \qquad (\hat{F}^{yy})_{ij} = \left(\frac{\partial\hat{\phi}_i}{\partial y}, \frac{\partial\psi_j}{\partial y}\right)_{\hat{n}\times m}.$$

Furthermore, we define $2n \times 2n$ matrices E_2, E_3 by

$$(E_2)_{ij} = \begin{pmatrix} (M^x\hat{F}^{xx} + M^y\hat{F}^{yx})G_b \\ (M^x\hat{F}^{xy} + M^y\hat{F}^{yy})G_b \end{pmatrix}_{2n\times 2n},$$

$$(E_3)_{ij} = \begin{pmatrix} -(K^x\hat{L}^{-1}(K^x)^{\mathrm{T}} + K^y\hat{L}^{-1}(K^y)^{\mathrm{T}})L^{-1} & 0 \\ 0 & -(K^x\hat{L}^{-1}(K^x)^{\mathrm{T}} - K^y\hat{L}^{-1}(K^y)^{\mathrm{T}})L^{-1} \end{pmatrix}_{2n\times 2n}.$$

Lemma 4.7. *We define the* $2n \times 2n$ *matrix* A_2 *by*

$$\begin{aligned}(A_2)_{ij} &= v^2(G_a)^{\mathrm{T}}E_1G_a - vG_aE_2 - v(G_aE_2)^{\mathrm{T}} - vG_aE_3 - v(G_aE_3)^{\mathrm{T}} \\ &\quad + G_b^{\mathrm{T}}EF + (G_b^{\mathrm{T}}EF)^{\mathrm{T}} + G_b^{\mathrm{T}}\tilde{D}G_b + F.\end{aligned}$$

Then, K_2 *can be estimated as follows:*

$$K_2 \leqslant \left(\sup_{\boldsymbol{x}\in\mathbb{R}^{2n}} \frac{\boldsymbol{x}^{\mathrm{T}}A_2\boldsymbol{x}}{\boldsymbol{x}^{\mathrm{T}}F\boldsymbol{x}}\right)^{1/2}. \tag{4.16}$$

Proof. Expanding $|v\overline{\Delta}u_h - \nabla p_h + P_0 f|_0^2$, we have

$$\begin{aligned}|v\overline{\Delta}u_h - \nabla p_h + P_0 f|_0^2 &= v^2(\overline{\Delta}u_h, \overline{\Delta}u_h) - v(\overline{\Delta}u_h, \nabla p_h) - v(\nabla p_h, \overline{\Delta}u_h) \\ &\quad + v(\overline{\Delta}u_h, P_0 f) + v(P_0 f, \overline{\Delta}u_h) - (\nabla p_h, P_0 f) \\ &\quad - (P_0 f, \nabla p_h) + (\nabla p_h, \nabla p_h) + |P_0 f|_0^2.\end{aligned}$$

We will represent each term by the quadratic forms of $2n$-dimensional vectors $\boldsymbol{f}$, which is obtained from the inner products of f as follows:

$$\begin{aligned}(\overline{\Delta}u_h, \overline{\Delta}u_h) &= \boldsymbol{a}E_1\boldsymbol{a}^{\mathrm{T}} \\ &= \boldsymbol{f}^{\mathrm{T}}(G_a)^{\mathrm{T}}E_1G_a\boldsymbol{f},\end{aligned}$$

$$\begin{aligned}(\overline{\Delta}u_h, \nabla p_h) &= \boldsymbol{a}_1\ (M^x\hat{F}^{xx} + M^y\hat{F}^{yx})\ \boldsymbol{b}^{\mathrm{T}} + \boldsymbol{a}_2\ (M^x\hat{F}^{xy} + M^y\hat{F}^{yy})\boldsymbol{b}^{\mathrm{T}} \\ &= (\boldsymbol{a}_1 \quad \boldsymbol{a}_2)\begin{pmatrix} (M^x\hat{F}^{xx} + M^y\hat{F}^{yx})G_b \\ (M^x\hat{F}^{xy} + M^y\hat{F}^{yy})G_b \end{pmatrix}\boldsymbol{f}\end{aligned}$$

$$
\begin{aligned}
&= \boldsymbol{a} E_2 \boldsymbol{f} \\
&= \boldsymbol{f}^{\mathrm{T}} G_a E_2 \boldsymbol{f},
\end{aligned}
$$

$$(\nabla p_h, \overline{\Delta} u_h) = \boldsymbol{f}^{\mathrm{T}} (G_a E_2)^{\mathrm{T}} \boldsymbol{f},$$

$$
\begin{aligned}
(\overline{\Delta} u_h, P_0 f) &= (\overline{\Delta} u_h^{(1)}, P_0 f_1) + (\overline{\Delta} u_h^{(2)}, P_0 f_2) \\
&= \boldsymbol{c}_1 (-K^x)^{\mathrm{T}} \boldsymbol{q}_1^{\mathrm{T}} + \boldsymbol{d}_1 (-K^y)^{\mathrm{T}} \boldsymbol{q}_1^{\mathrm{T}} + \boldsymbol{c}_2 (-K^x)^{\mathrm{T}} \boldsymbol{q}_2^{\mathrm{T}} + \boldsymbol{d}_2 (-K^y)^{\mathrm{T}} \boldsymbol{q}_2^{\mathrm{T}} \\
&= (\boldsymbol{a}_1 \;\; \boldsymbol{a}_2) \begin{pmatrix} -(M^x (K^x)^{\mathrm{T}} + M^y (K^y)^{\mathrm{T}}) L^{-1} & 0 \\ 0 & (M^x (-K^x)^{\mathrm{T}} + M^y (-K^y)^{\mathrm{T}}) L^{-1} \end{pmatrix} \begin{pmatrix} \boldsymbol{f}_1 \\ \boldsymbol{f}_2 \end{pmatrix} \\
&= \boldsymbol{a} \begin{pmatrix} -(K^x \hat{L}^{-1} (K^x)^{\mathrm{T}} + K^y \hat{L}^{-1} (K^y)^{\mathrm{T}}) L^{-1} & 0 \\ 0 & -(K^x \hat{L}^{-1} (K^x)^{\mathrm{T}} - K^y \hat{L}^{-1} (K^y)^{\mathrm{T}}) L^{-1} \end{pmatrix} \boldsymbol{f} \\
&= \boldsymbol{f}^{\mathrm{T}} G_a E_3 \boldsymbol{f},
\end{aligned}
$$

$$(P_0 f, \overline{\Delta} u_h) = \boldsymbol{f}^{\mathrm{T}} (G_a E_3)^{\mathrm{T}} \boldsymbol{f},$$

$$
\begin{aligned}
(\nabla p_h, P_0 f) &= \left(\frac{\partial p_h}{\partial x}, P_0 f_1 \right) + \left(\frac{\partial p_h}{\partial y}, P_0 f_2 \right) \\
&= \boldsymbol{b} (-E_x \boldsymbol{q}_1^{\mathrm{T}} - E_y \boldsymbol{q}_2^{\mathrm{T}}) \\
&= \boldsymbol{b} (-E_x L^{-1} \boldsymbol{f}_1 - E_y L^{-1} \boldsymbol{f}_2) \\
&= -\boldsymbol{b} (E_x \;\; E_y) \begin{pmatrix} L^{-1} & 0 \\ 0 & L^{-1} \end{pmatrix} \begin{pmatrix} \boldsymbol{f}_1 \\ \boldsymbol{f}_2 \end{pmatrix} \\
&= -\boldsymbol{f}^{\mathrm{T}} G_b^{\mathrm{T}} E F \boldsymbol{f},
\end{aligned}
$$

$$(P_0 f, \nabla p_h) = -\boldsymbol{f}^{\mathrm{T}} (G_b^{\mathrm{T}} E F)^{\mathrm{T}} \boldsymbol{f},$$

$$
\begin{aligned}
(\nabla p_h, \nabla p_h) &= \boldsymbol{b} \tilde{D} \boldsymbol{b}^{\mathrm{T}} \\
&= \boldsymbol{f}^{\mathrm{T}} G_b^{\mathrm{T}} \tilde{D} G_b \boldsymbol{f}.
\end{aligned}
$$

Therefore, we have

$$|\nu \overline{\Delta} u_h - \nabla p_h + P_0 f|_0^2 = \boldsymbol{f}^{\mathrm{T}} A_2 \boldsymbol{f}. \qquad \square$$

Next, we define $n \times n$ matrices D^{xx}, D^{xy} and D^{yy} as

$$(D^{xx})_{ij} = \left(\frac{\partial \phi_i}{\partial x}, \frac{\partial \phi_j}{\partial x} \right)_{n \times n}, \qquad (D^{xy})_{ij} = \left(\frac{\partial \phi_i}{\partial x}, \frac{\partial \phi_j}{\partial y} \right)_{n \times n}, \qquad (D^{yy})_{ij} = \left(\frac{\partial \phi_i}{\partial y}, \frac{\partial \phi_j}{\partial y} \right)_{n \times n},$$

and $2n \times 2n$ matrix Q_3 as

$$(Q_3)_{ij} = \begin{pmatrix} D^{xx} & D^{xy} \\ (D^{xy})^{\mathrm{T}} & D^{yy} \end{pmatrix}_{2n \times 2n}.$$

Then, we obtain the following lemma using these matrices.

Lemma 4.8. *We define the* $2n \times 2n$ *matrix* A_3 *by*

$$(A_3)_{ij} = (G_a Q_3 G_a)_{2n\times 2n}.$$

Then, K_3 *can be estimated as follows:*

$$K_3 \leqslant \left(\sup_{x\in\mathbb{R}^{2n}} \frac{\boldsymbol{x}^{\mathrm{T}} A_3 \boldsymbol{x}}{\boldsymbol{x}^{\mathrm{T}} F \boldsymbol{x}} \right)^{1/2}. \tag{4.17}$$

Proof. From (4.12), we have

$$\begin{aligned} |\operatorname{div} u_h|_0^2 &= \boldsymbol{a}_1 D^{xx} \boldsymbol{a}_1^{\mathrm{T}} + \boldsymbol{a}_1 D^{xy} \boldsymbol{a}_2^{\mathrm{T}} + \boldsymbol{a}_2 (D^{xy})^{\mathrm{T}} \boldsymbol{a}_1^{\mathrm{T}} + \boldsymbol{a}_2 D^{yy} \boldsymbol{a}_2^{\mathrm{T}} \\ &= \boldsymbol{a} Q_3 \boldsymbol{a}^{\mathrm{T}} \\ &= \boldsymbol{f}^{\mathrm{T}} G_a Q_3 G_a \boldsymbol{f} \\ &= \boldsymbol{f}^{\mathrm{T}} A_3 \boldsymbol{f}. \qquad \square \end{aligned}$$

Finally, we obtain the following estimate of K_4.

Lemma 4.9. *We define the* $2n \times 2n$ *matrix* A_4 *by*

$$(A_4)_{ij} = (G_b^{\mathrm{T}} EF + (G_b^{\mathrm{T}} EF)^{\mathrm{T}} + G_b^{\mathrm{T}} \tilde{D} G_b + F)_{2n\times 2n}.$$

Then, K_4 *can be estimated as follows:*

$$K_4 \leqslant \left(\sup_{x\in\mathbb{R}^{2n}} \frac{\boldsymbol{x}^{\mathrm{T}} A_4 \boldsymbol{x}}{\boldsymbol{x}^{\mathrm{T}} F \boldsymbol{x}} \right)^{1/2}. \tag{4.18}$$

Proof. Expanding $|-\nabla p_h + P_0 f|_0^2$, by the proof of Lemma 4.7, we obtain

$$\begin{aligned} |-\nabla p_h + P_0 f|_0^2 &= (\nabla p_h, \nabla p_h) - (\nabla p_h, P_0 f) - (P_0 f, \nabla p_h) + |P_0 f|_0^2 \\ &= \boldsymbol{f}^{\mathrm{T}} G_b^{\mathrm{T}} EF \boldsymbol{f} + \boldsymbol{f}^{\mathrm{T}} (G_b^{\mathrm{T}} EF)^{\mathrm{T}} \boldsymbol{f} + \boldsymbol{f}^{\mathrm{T}} G_b^{\mathrm{T}} \tilde{D} G_b \boldsymbol{f} + \boldsymbol{f}^{\mathrm{T}} F \boldsymbol{f} \\ &= \boldsymbol{f}^{\mathrm{T}} (G_b^{\mathrm{T}} EF + (G_b^{\mathrm{T}} EF)^{\mathrm{T}} + G_b^{\mathrm{T}} \tilde{D} G_b + F) \boldsymbol{f} \\ &= \boldsymbol{f}^{\mathrm{T}} A_4 \boldsymbol{f}. \qquad \square \end{aligned}$$

Note that estimates (4.15)–(4.18) are reduced to finding the maximum eigenvalue of

$$Ax = \lambda Bx$$

where A is a $2n \times 2n$ symmetric matrix, B a $2n \times 2n$ symmetric and positive definite matrix, respectively. Then, using a procedure proposed by [15], we can estimate these eigenvalues.

5. Numerical examples

In this section we give several numerical examples for a posteriori and a priori error bounds.

Table 1

N	$\|u-u_h\|_1$	$\|p-p_h\|_0$
5	1.39353	4.70242
10	0.34405	1.16099
15	0.15184	0.51237
20	0.08502	0.28689
25	0.05421	0.18292
30	0.03752	0.12660

Let Ω be a rectangular domain in $\mathbb{R}^2$ such that $\Omega=(0,1)\times(0,1)$. We consider the following Stokes problem:

$$\begin{aligned} -\Delta u+\nabla p&=f \quad \text{in}\,\Omega=(0,1)\times(0,1),\\ \operatorname{div} u&=0 \quad \text{in}\,\Omega.\\ u&=0 \quad \text{on}\,\partial\Omega. \end{aligned} \tag{5.1}$$

Also let δ_x: $0=x_0<x_1<\cdots<x_N=1$ be a uniform partition, and let δ_y be the same partition as δ_x for y direction. We define the partition of Ω by $\delta\equiv\delta_x\otimes\delta_y$. N denotes the number of partitions for the interval $(0,1)$, i.e. $h=1/N$.

Further, we define the finite element subspace X_h and Y_h by $X_h\equiv\mathcal{M}_0^2(x)\otimes\mathcal{M}_0^2(y)$ where $\mathcal{M}_0^2(x)$, $\mathcal{M}_0^2(y)$ are sets of continuous piecewise quadratic polynomials on $(0,1)$ under the above partition δ with homogeneous boundary condition and set $Y_h\equiv\mathcal{M}_0^1(x)\otimes\mathcal{M}_0^1(y)\cap L_0^2(\Omega)$ where $\mathcal{M}_0^1(x)$, $\mathcal{M}_0^1(y)$ piecewise linear as well. Then, the matrix G in (4.11) is invertible because the space $X_h^2\times Y_h$ satisfies the usual discrete inf-sup condition [4].

We can also take the constant $v=1$, $C_0=1/(2\pi)$ [8] and $1/\beta^2=4+2\sqrt{2}$.

5.1. A posteriori error bounds

We take the vector function $f=(f_1,f_2)^{\mathrm{T}}$ as

$$f_1=50(-2x+y+xy),$$
$$f_2=20(1-5xy).$$

In this case, $\|u_h\|_{L^\infty(\Omega)}\sim 0.52$ and $\|p_h\|_{L^\infty(\Omega)}\sim 22.72$, where $\|\cdot\|_{L^\infty(\Omega)}$ is the L^∞-norm on Ω. We obtained a posteriori error bounds $|u-u_h|_1$ and $|p-p_h|_0$ given in Table 1.

Figs. 1 and 2 shows the pressure and vector field on Ω, respectively.

Next, in order to examine the quality of our estimator, we choose the vector function f so that $u=(u_1,u_2)^{\mathrm{T}}$:

$$u_1(x,y)=20x^2(1-x)^2y(1-y)(1-2y),$$
$$u_2(x,y)=-20y^2(1-y)^2x(1-x)(1-2x),$$
$$p(x,y)=4x(-1+2y)(10x^2-15x^3+6x^4-10y+30xy-20x^2y+10y^2-30xy^2+20x^2y^2)$$

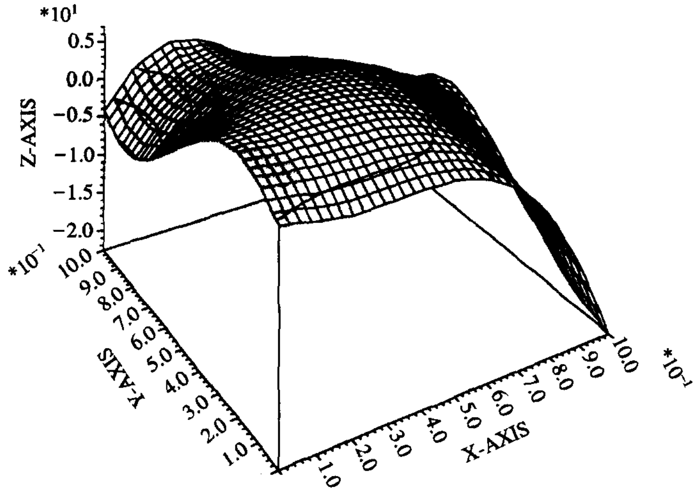

Fig. 1. Pressure field p.

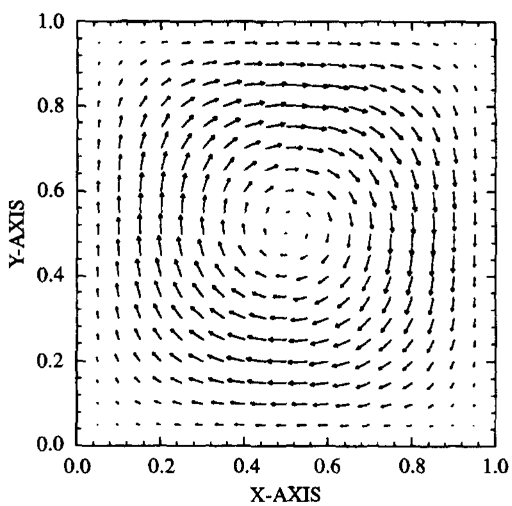

Fig. 2. Vector field $u = (u_1, u_2)$.

are the exact solutions for (1.1). In this case, $|u|_1 = 4/7 \sim 0.571$ and $|p|_0 = 2\sqrt{962/33}/7 \sim 1.543$. We should adopt the quantity $C(u_h, p_h)$ as the error estimator for both of the velocity and pressure instead of the right-hand side of (3.10), because the error indicator is usually evaluated by the quantity exclusive of the proportional constant independent of mesh size h. We obtained each relative a posteriori error bounds $|u - u_h|_1/|u|_1$, $|p - p_h|_0/|p|_0$ given in Table 2.

In this case, we can compute exact norms $|u - u_h|_1$ and $|p - p_h|_0$. The ratio of the relative errors between a posteriori error estimator by Theorem 3.1 to the exact norms is nearly independent of N, namely 1.4 for the velocity and 1.3 for the pressure.

These examples show that our *mathematically rigorous* a posteriori error bounds have rate of convergence with *optimal* order even if the exact solutions are unknown.

Table 2

N	$\lvert u - u_h\rvert_1/\lvert u\rvert_1$	$\lvert p - p_h\rvert_0/\lvert p\rvert_0$
5	0.37895	0.47369
10	0.08976	0.11220
15	0.03893	0.04866
20	0.02159	0.02698
25	0.01369	0.01711
30	0.00943	0.01180

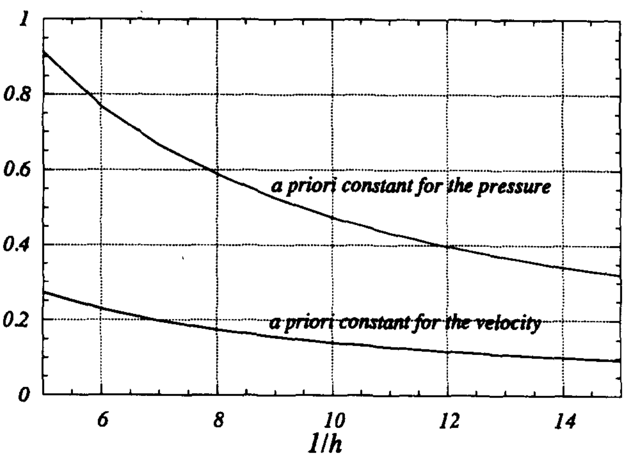

Fig. 3. A priori error constants.

5.2. Constructive a priori error constants

We computed the a priori constants by using Theorem 4.3 because the computation is much simpler than by Theorem 4.1. Fig. 3 illustrates a priori error constants for the velocity and pressure of (4.9):

$$\left(\frac{1}{\nu^2}+\frac{1}{\beta^2}\right)^{1/2} C(h) \quad \text{and} \quad \left(\frac{1}{\beta}+\frac{\nu}{\beta^2}\right) C(h),$$

respectively. These results confirm us the expected rate of convergence with optimal order.

The numerical examples are computed on FUJITSU VP2600/10 vector processor by the usual computer arithmetic with double precision. So, the round off errors in these examples are neglected. However, it should be sufficient for our present purposes.

References

[1] R. E. Bank, B. D. Welfert, A posteriori error estimates for the Stokes equations: a comparison, Comput. Meth. Appl. Mech. Eng. 82 (1990) 323–340.

[2] R. E. Bank, B. D. Welfert, A posteriori error estimates for the Stokes problem, SIAM J. Numer. Anal. 28 (1991) 591–623.

[3] P. G. Ciarlet, The Finite Element Method for Elliptic Problems, North-Holland, Amsterdam, 1978.
[4] V. Girault, P. A. Raviart, Finite Element Approximation of the Navier–Stokes equations, Series in Computational Mathematics, Springer, Berlin, 1986.
[5] C. O. Horgan, L. E. Payne, On inequalities of Korn, Friedrichs and Babuška-Aziz, Arch. Rat. Mech. Anal. 82 (1983) 165–179.
[6] M. T. Nakao, A numerical approach to the proof of existence of solutions for elliptic problems, Japan J. Appl. Math. 5 (1988) 313–332.
[7] M. T. Nakao, Solving nonlinear elliptic problems with result verification using an H^{-1} type residual iteration, Computing, Suppl. 9 (1993) 161–173.
[8] M. T. Nakao, N. Yamamoto, S. Kimura, On best constant in the optimal error estimates for the H_0^1-projection into Piecewise Polynomial Spaces. J. Approx. Theory, to appear.
[9] A. Russo, A posteriori error estimators for the Stokes problem, Appl. Math. Lett. 8 (1995) 1–4.
[10] M. H. Schultz, Spline Analysis, Prentice-Hall, Englewood Cliffs, NJ, 1973.
[11] R. Verfürth, A posteriori error estimators for the Stokes equations, Numer. Math. 55 (1989) 309–325.
[12] R. Verfürth, A posteriori error estimators for the Stokes equations II non-conforming discretizations, Numer. Math. 60 (1991) 235–249.
[13] R. Verfürth, A posteriori error estimates for nonlinear problems. Finite element discretizations of elliptic equations, Math. Comp. 62 (1994) 445–475.
[14] Y. Watanabe, M. T. Nakao, Numerical verifications of solutions for nonlinear elliptic equations, Japan J. Ind. Appl. Math. 10 (1993) 165–178.
[15] N. Yamamoto, M. T. Nakao, Numerical verifications of solutions for elliptic equations in nonconvex polygonal domains, Numer. Math. 65 (1993) 503–521.
[16] N. Yamamoto, M. T. Nakao, Numerical verifications for solutions to elliptic equations using residual iterations with a higher order finite element, J. Comput. Appl. Math. 60 (1995) 271–279.